Defense Technical Information Center Compilation Part Notice

ADP013500

TITLE: Issues in the Design and Optimization of Health Management Systems

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: New Frontiers in Integrated Diagnostics and Prognostics. Proceedings of the 55th Meeting of the Society for Machinery Failure Prevention Technology. Virginia Beach, Virginia, April 2 - 5, 2001

To order the complete compilation report, use: ADA412395

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP013477 thru ADP013516

ISSUES IN THE DESIGN AND OPTIMIZATION OF HEALTH MANAGEMENT SYSTEMS

Michael Yukish, Carl Byington, and Robert Campbell

The Pennsylvania State University
Applied Research Laboratory
P.O. Box 30 (North Atherton Street)
State College, Pennsylvania 16804-0030

Abstract: The design of a health management system is presented as a decision problem. The decision space is affected by the choice of a particular health management system design and an employed maintenance policy. To the 1st order, the evaluation objectives consist of the conflicting goals of minimizing purchase costs, minimizing operating costs, and maximizing availability. In order to assist and even automate the decision process, data and computational tools for calculating the objectives are needed. While calculating purchase costs is straightforward, determining operating costs and availability is not. Parameters such as failure rate, criticality, component replacement cost due to unplanned and planned maintenance, and average downtime for repair are examples of data needed to determine the operating costs and availability. These types of data are not part of traditional product models. Some of the data is partially contained in the traditional FMECA, but much of it is not. This shortcoming is the motivation for tools to assist in the design of health management systems, such as the FMECA++© tool being developed by Impact Technologies and Penn State Applied Research Laboratory.

Key Words: Availability; CBM; evaluation metrics; FMECA; health management; operating cost; optimization; purchase cost.

Introduction: It is well known that health management systems can increase the overall reliability of the underlying system by providing early fault detection and diagnostic localization. In the ultimate case, a CBM system would enable one to predict the remaining useful life of critical components, and to isolate the root cause of failures after the failure symptoms have been observed. If predictions can be made, replacement part orders and repair actions can be optimally scheduled to reduce the overall operational and maintenance-related costs, while minimizing downtime and therefore maximizing system availability. These improvements in operating costs and availability are of course offset by the increase in cost of acquiring and maintaining the health management system.

Thus, the choice of what health management system to use can be abstractly considered as a decision problem, where the decision maker chooses a health management system and an accompanying maintenance policy to satisfy the conflicting goals of minimizing purchase costs, minimizing operating costs, and maximizing availability. Cast in this

fashion, the tools and techniques of multi-objective optimization and multidisciplinary design optimization can be used to find a "best" design [1].

In order to assist and even automate the decision process, data and computational tools for calculating the objectives are needed. While calculating purchase costs is straightforward, determining operating costs and availability is not. Parameters such as failure rate, criticality, component replacement cost due to unplanned and planned maintenance, and average downtime for repair are examples of data needed to determine the operating costs and availability. Again, such data is partially contained in the traditional FMECA, but much of it is not. Augmented models such as used in the FMECA++© intrinsically capture the downstream effects of the failure modes, including secondary effects as embodied in a hierarchical model. FMECA++© is envisioned to be a (graphical & tabular) representation of functional failure modes with hierarchically linked effects and symptoms to provide a blueprint for the design of a health management system. It extends a typical FMECA with information on precursor symptoms, sensor observables, diagnostic/prognostic processes and their associated metrics. The data embodied in the FMECA++© can be combined with its associated methods and tools for calculating operating costs and availability, and the problem can be cast as a multi-objective optimization problem and solved using well-known methods.

The remainder of this paper first establishes the basic problem statement for casting the choice of a health management system as a decision problem, choosing over multiple objectives. Next various methods for optimizing with multiple objectives are presented. The determination of an availability metric receives extra attention, as its calculation is less straightforward in comparison to the other objectives. Finally, the requirements imposed on a design environment in order to implement the problem structure developed in this paper are presented.

Statement of the Decision Problem: In choosing a health management system, the decision maker starts (by assumption) with a system to be monitored (S), and has the conflicting objectives of minimizing purchase cost (PC) and operating cost (OC) while maximizing availability (A). The "decision space" is the choice of health monitoring suite to employ (HM), and the choice of accompanying maintenance policy (MP). The prime dependencies of the objectives with regard to the decisions are as follows:

$$\begin{aligned} \text{Purchase Cost} &= \mathrm{PC}(\mathrm{S}, \mathrm{HM}) \\ \text{Operating Cost} &= \mathrm{OC}(\mathrm{S}, \mathrm{MP}, \mathrm{HM}) \\ \text{Availability} &= \mathrm{A}(\mathrm{S}, \mathrm{MP}, \mathrm{HM}) \end{aligned} \tag{1}$$

Note that in addition to the dependence on the system to be monitored, purchase cost is shown as a function of the choice of health management system, and operating cost is shown as a function of the maintenance policy and the heath management system. Particularly in the case of operating costs, this is done to make explicit the dependency of the operating cost on both the maintenance policy and the health management system. The health management system will directly affect the operating costs to a small degree through its own life cycle costs. It will also affect OC with the ability to impact the required amount of maintenance and provide potentially large mishap cost avoidances.

Multi-Objective Optimization: The health management choice presents a multi-objective decision problem, where the objectives are conflicting. In this instance, a significant tradeoff is in the up-front purchase cost of a health management system versus the downstream savings in operating costs. An additional tradeoff is, given a health management system, choosing a maintenance policy that will minimize the operating costs versus choosing one to maximize the availability. Many methods are available for finding "best" solutions for such problems [2]. Each method attempts to capture, in some rational manner, the decision maker's preference. The methods discussed briefly here are weighted sums, minimax, goal programming, and *design by shopping*, explained below.

The most basic methods are weighted sums methods, where a scalar measure of worth is calculated by multiplying each of PC, OC and A by a weighting factor. Note that the availability term is subtracted from the total, to account for its maximization vice the other terms' minimization:

$$\begin{gathered} \min_{\mathrm{HM,MP}} \quad w_1\mathrm{PC} + w_2\mathrm{OC} - w_3\mathrm{A} \\ \text{where } \sum_{i=1}^{3} w_i = 1 \end{gathered} \tag{2}$$

The weighting parameters are an attempt to capture the preference of the decision maker as to the relative importance of the terms. These can be generalized to quadratic and higher systems with weighted sums:

$$\begin{gathered} \min_{\mathrm{HM,MP}} \quad w_1\mathrm{PC}^k + w_2\mathrm{OC}^k - w_3\mathrm{A}^k \\ \text{where } \sum_{i=1}^{3} w_i = 1, \quad k \in \{1,2,3,\ldots\} \end{gathered} \tag{3}$$

Another method is to apply the minimax criteria as follows:

$$\begin{gathered} \min_{\mathrm{HM,MP}} \quad \max_{w_i}\left[w_1\mathrm{PC} + w_2\mathrm{OC} - w_3\mathrm{A}\right] \\ \text{where } \sum_{i=1}^{3} w_i = 1 \end{gathered} \tag{4}$$

The minimax criteria can be interpreted as "chose HM and MP so as to minimize the worst possible choice of weights *w*." Use of the minimax criteria can be construed as an attempt to guard against incorrectly capturing a user's preference, expressed in *w*. However, designs chosen by the minimax criteria are usually considered as too conservative.

Another method is known as pre-emptive goal programming. With this method, the objectives are first ordered from most to least important. Then the optimization problem is solved for the most important objective first, and only solving for the next objective if the answer to the first problem is non-unique. So for example if the objectives are ordered {PC, OC, A}, then the problem solved first is

$$PC^{*} = \min_{HM,MP} \quad PC(HM, MP) \tag{5}$$

If the solution for HM is not unique, then the next problem solved is

$$\begin{aligned} OC^{*} &= \min_{HM,MP} \quad OC(HM,MP) \\ \text{s.t.} \quad & PC(HM, MP) = PC^{*} \end{aligned} \tag{6}$$

and so on until a unique solution is reached.

The final method presented, known as design by shopping, does not establish a global objective at all [3]. Rather, the *pareto frontier* of the feasible results of PC, OC, and A is presented to the user, and the user decides. The pareto frontier is the set of all designs that are *non-dominated* by other designs. Assume that a choice of HM and MP result in a set of values {PC, OC, A} that define a point in the performance space. This point is non-dominated if an improvement in any one of the objectives can only be achieved by a decrease in one or more of the others. A dominated design point is one where a feasible design exists that is at least as good as the first point in all objectives, and better in at least one. The diagram below, Figure 1, shows the pareto frontier in bold for the two-dimensional case, holding purchase cost constant. Designs that are along the lower left boundary are non-dominated, in that an improvement in one aspect is accompanied by a decrement in another. Interior points are dominated, and it is reasonable to expect that a decision maker would never choose one.

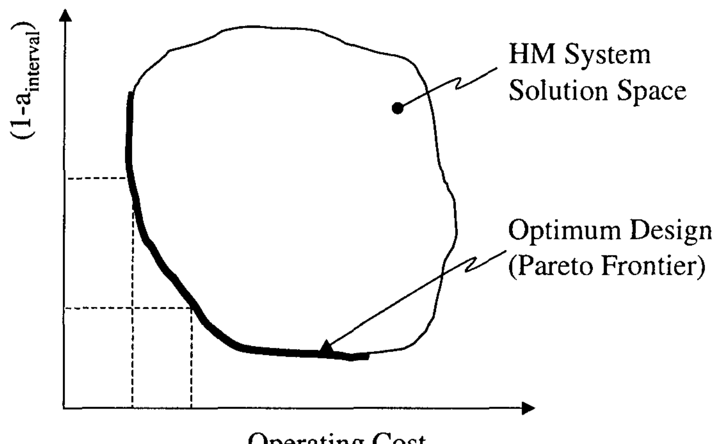

Figure 1: Pareto Frontier

Many design optimization experts would now consider it a mistake to come up with some single scalar objective that is a blend of PC, OC and A and that attempts to capture the preference of the decision maker. In their opinion it is better to let the decision maker

"shop" for the right mix by presenting designs from the pareto set, rather than trying to capture the decision maker's preference, which experience has proven difficult.

Determining Purchasing and Operating Cost: Purchasing cost can be determined using parametric cost estimating applications such as PRICE-E ™. Such tools have been widely employed in the cost estimating of conceptual through detailed design [4]. The primary inputs are weight, volume and complexity of the subsystems, the hierarchical structure of the system, and the complexity of the assemblies. If the baseline costs are known due to activity-based costing, then these numbers can be used. Additional data needed to estimate purchasing costs relates to the actual purchase, e.g., the dates for initiating purchases, buy rates, total amount purchased, and so on. The same tools can also be used to develop approximations to operating costs, based on the design data listed.

Determining availability: It is probably too hard to calculate availability in closed form for a system that forms a complex reliability block diagram. Upper and lower bounds might be calculated by making simplifying assumptions, such as choosing failure rate statistics from a restrictive set of families, or assuming the reliability block diagram is all series or parallel. But if models of the system are available, tools exist to simulate the system and determine an availability metric. A choice is to use a Monte Carlo simulation, with values for isolation, repair, and admin/logistics times for various components along with failure statistics, and simulate the system over an interval to get an estimate of availability.

Availability metric: Presented in this section is one approach to determining a measure of availability. It is important to bear in mind that, if availability is to be an objective, its computation must be such that the impact of the addition of HM is clear.

For a system that operates over some fixed interval, the availability can be determined by the equation

$$a_{\text{interval}} = \frac{T_{\text{operate}}}{T_{\text{iso}} + T_{\text{adl}} + T_{\text{repair}} + T_{\text{PM}} + T_{\text{operate}}} \tag{7}$$

where over the interval, T_{iso} is the time spent isolating faults, T_{adl} is the admin and logistics time associated with repair events, T_{repair} is the time to disassemble, repair or replace, and reassemble during repair events, T_{PM} is the time spent doing preventive maintenance, and T_{operate} is the time operating [5]. All of the T's other than T_{operate} must have the caveat that they only count if they occurred when the system was supposed to be available. So the additional constraint can be imposed

$$T_{\text{required}} = T_{\text{iso}} + T_{\text{adl}} + T_{\text{repair}} + T_{\text{PM}} + T_{\text{operate}} \tag{8}$$

where T_{required} is the time that the system is required to be operating, which may be only eight hours per day, for example. In this case, all maintenance may be performed while the system is not required to be available, ensuring 100% availability.

To gain a feel for how the addition of health management can affect the various parameters, we consider below the results of adding a partially effective and a *perfect* health management suite compared to no health monitoring [Table 1]. Reference [10] discusses previously developed metrics to associate with a HM system design, and reference [9] proposes how this could be applied to produce an operational impact on A and OC. These metrics represent a way to propagate the effectiveness of specific HM sensors and algorithms and map them to an availability effect as is shown in Table 1. To further simplify the problem of comparison, assume the available maintenance actions are restricted to preventive maintenance (PM) and replacement (REP). Further assume that a REP is either planned or unplanned.

Table 1: Comparison with degrees of Health Management

	No HM	**Partially Effective HM**	**Perfect HM**
Unplanned replacement	$T_{\text{repair}}, T_{\text{iso}}, T_{\text{adl}}$	$T_{\text{repair}}, T_{\text{iso}}, T_{\text{adl}}$. Unplanned replacements reduced corresponding to HM fault detection metrics	Unplanned replacements are totally eliminated
Planned replacement	$T_{\text{repair}}, T_{\text{iso}}$	$T_{\text{repair}}, T_{\text{iso}}$. Isolation time is reduced appropriately by diagnostic accuracy metric	T_{repair} only. Isolation time is eliminated
Preventive maintenance	T_{PM}	T_{PM} will be reduced, based upon the composite effectiveness of the HM system to predict the overall failure modes	T_{PM} will be reduced to provide predictive maintenance on all critical systems

Maintenance Policy: Complicating the decision problem greatly is the fact that the choice of maintenance policy has a critical impact on the value of the T variables. They can all be written as $T = T(\text{MP})$. For example, a HM system coupled with a maintenance policy that reads "Replace only on failure" will show no benefits of a health management system. In general, the choice of maintenance policy (MP) will have an impact on availability equal or greater in scope to the choice of HM system. Restating the equation for determining availability, (1)

$$\text{A} = \text{A}(\text{S}, \text{HM}, \text{MP}) \qquad (9)$$

If we are trying to find the HM system that gives us the best availability, we must solve the optimization problem for maximum availability while including MP as a decision variable.

$$A^* = \max_{HM,\, MP} \quad A(S, HM, MP) \tag{10}$$

Casting as a search for the best HM and moving the optimization with regard to MP to an inner loop, the problem gains a bi-level optimization structure.

$$HM^* = \arg\max_{HM} \left[\max_{MP} \quad A(S, HM, MP) \right] \tag{11}$$

Notional Infrastructure for Supporting the Decision Process: Given the statement of the decision problem above, the suite of computational tools and applications must next be developed. One such structure for supporting the decision process is shown below, in Figure 2.

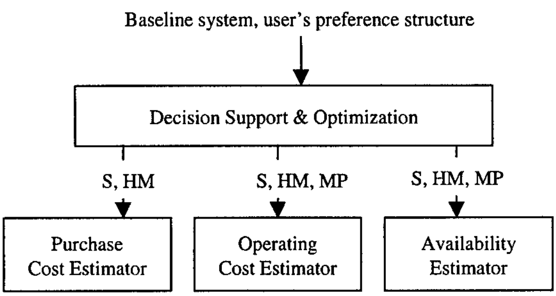

Figure 2: Decision Support Structure

At the top, the baseline system that is to be considered along with possibly some preference structure is entered. At the bottom are three separate applications, each of which analyzes a design concept to determine its value with respect to one of the three objectives. In the middle is the optimization engine, which in effect automates the search through the design space in order to find the "best" designs.

It is important to note that each of the three estimators require data about the system, both the baseline system and the chosen health management system, to be passed down, but that the constitution of the data differs from one estimator to another. The purchase cost estimator needs sizes, weights, complexities, and other manufacturing cost-related data of the system and its components. The availability estimator needs data about the components of the system, such as failure statistics as a function of loading, and about the constitution of the coupling of the components in the system, such as captured in a reliability block diagram. Therefore, before any optimization can occur, the product must be modeled in a fashion that can serve as input to the estimators.

Optimization Methods: Once having posed the decision problem and developed the appropriate data models to drive the estimators, the implementation of an optimization algorithm can be considered. The choice of an optimization algorithm is constrained by a number of aspects of the problem. First, the estimator inputs will likely contain a mix of continuous, countable, and enumerated variables. This implies a smooth optimization algorithm will not suffice for the overall problem, but may be applicable for sub-problems. However, if the problem is cast such that the maintenance policy is solved for in an inner loop and the health management choice is solved for in an outer loop, this presents a bi-level optimization problem. Bi-level optimization problems are notoriously difficult for gradient-based optimizers to work with, [6, 7].

Alternatives to the gradient-based optimization algorithms are the non-gradient methods such as simulated annealing and genetic algorithms. Genetic algorithms have the added benefit that they are conducive to exploring the pareto set of a design space, [8]. At each iteration, a new set of proposed designs are created, and the non-dominated designs are culled from the offspring. Eventually the genetic algorithm will develop a set of design points that are reasonably expected to be along the pareto set. A drawback to all of the non-gradient based methods is that they have no obvious stopping criteria, as does exist in the gradient-based methods.

Future Work: Because of its potential impact, health management solutions should be considered during the initial design of a system. However, current practice in system design does not adequately support the consideration of such solutions. It would appear that, because an initial system FMECA is performed during the design stage, it is a perfect link to the critical overall system failure modes that a health management system is designed to help mitigate. In fact, a process has been demonstrated that links this traditional FMECA analysis with health management system design optimization based on failure mode coverage and availability and life cycle cost analyses, [9]. But in order to be able to truly evaluate the relative merits of different health management system options, the systems must be modeled in a more extensive manner. New tools such as the FMECA++© are now being developed to address this shortfall, [9]. The methods presented herein can be implemented in such tools for use in the optimization of the system and the HM, thus providing the maximum benefit of HM through its impact on the system design.

Acknowledgment: This work was supported by a Navy Small Business Innovative Research Program through a subcontract with Impact Technologies (Contract No. S00139N). The content of the information does not necessarily reflect the position or policy of the Government, and no official endorsement should be inferred.

1. Sobieszczanski-Sobieski, J. *Multidisciplinary Design Optimization: An Emerging New Engineering Discipline.* in *The World Congress on Optimal Design of Structural Systems.* 1993. Rio de Janeiro, Brazil: NASA Technical Memorandum 107761.

2. Rosenthal, R., *Principles of Multiobjective Optimization.* Decision Sciences, 1985. **16**: p. 133-152.
3. Balling, R. *Design by Shopping: A New Paradigm?* in *Proceedings of the Third World Congress of Structural and Multidisciplinary Optimization (WCSMO-3).* 1999. Buffalo, NY: University at Buffalo.
4. Yukish, M., L. Bennett, and T. Simpson. *Requirements on MDO imposed by the Undersea Vehicle Conceptual Design Problem.* in *8th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization.* 2000. Long Beach, CA: AIAA.
5. Dhillon, B.S., *Engineering Maintainability.* 1999, Houston, TX: Gulf Publishing Company.
6. Vicente, L. and P. Calamai, *Bilevel and multilevel programming: a bibliography review.* 2000.
7. Alexandrov, N.M. and R. Lewis, *Analytical and computational aspects of collaborative optimization.* 2000, NASA: Hampton, VA. p. 26.
8. Murata, T. and H. Ishibuchi. *MOGA: multi-objective genetic algorithms.* in *IEEE Conference on Evolutionary Computation.* 1995: IEEE.
9. Kacprzynski, G., et al. *Enhanced FMECA: Integrating Health Management Design and Traditional Failure Analysis.* in *55th Meeting of the MFPT.* 2001. Virginia Beach, Virginia.
10. Orsagh, R., et al., *Development of Metrics for Mechanical Diagnostic Technique Qualification and Validation,* Proceedings of the COMADEM, Houston TX, December 2000.